

Gephi and Network Analysis

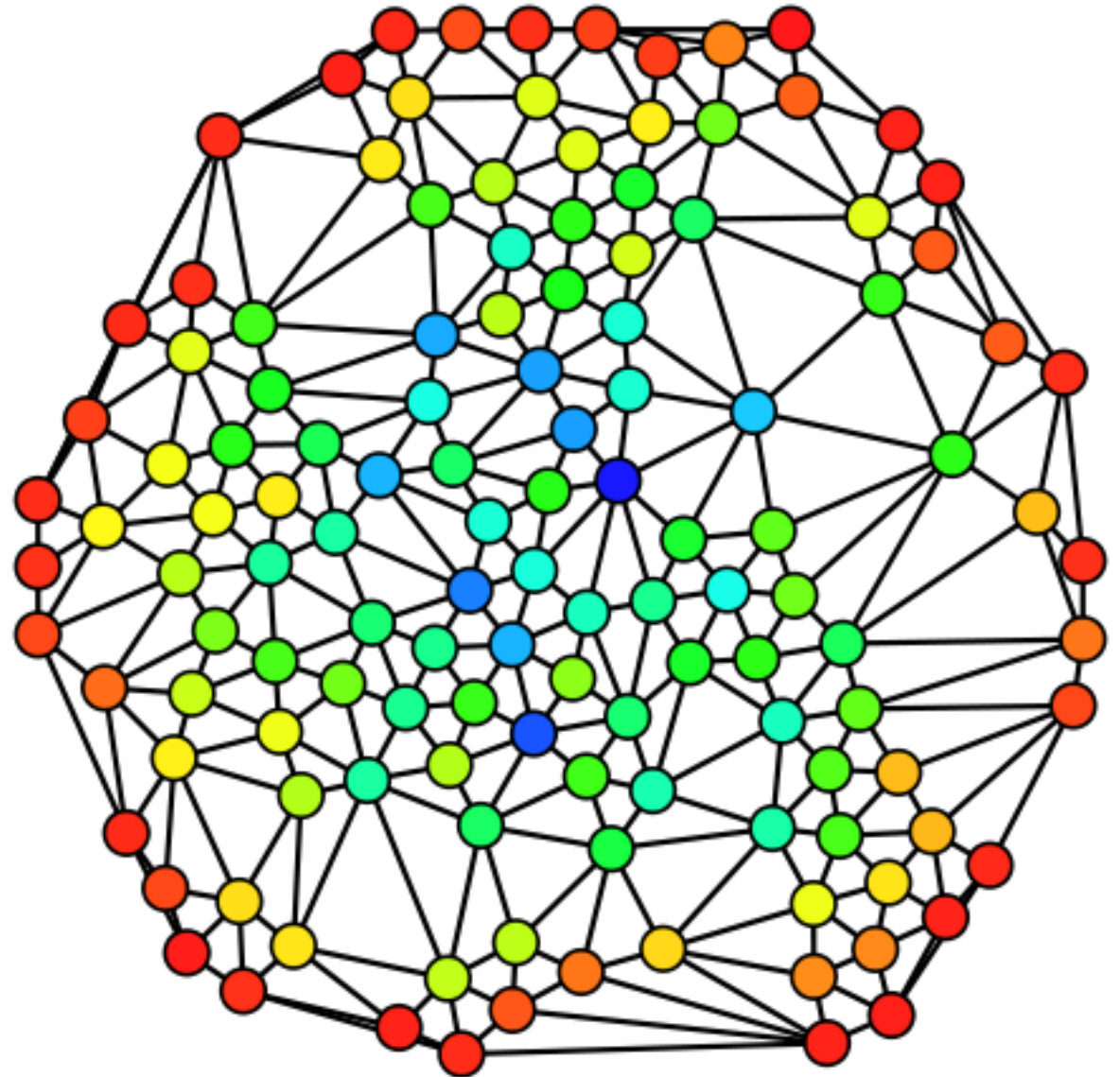
Jen Stevens

Hack To Learn

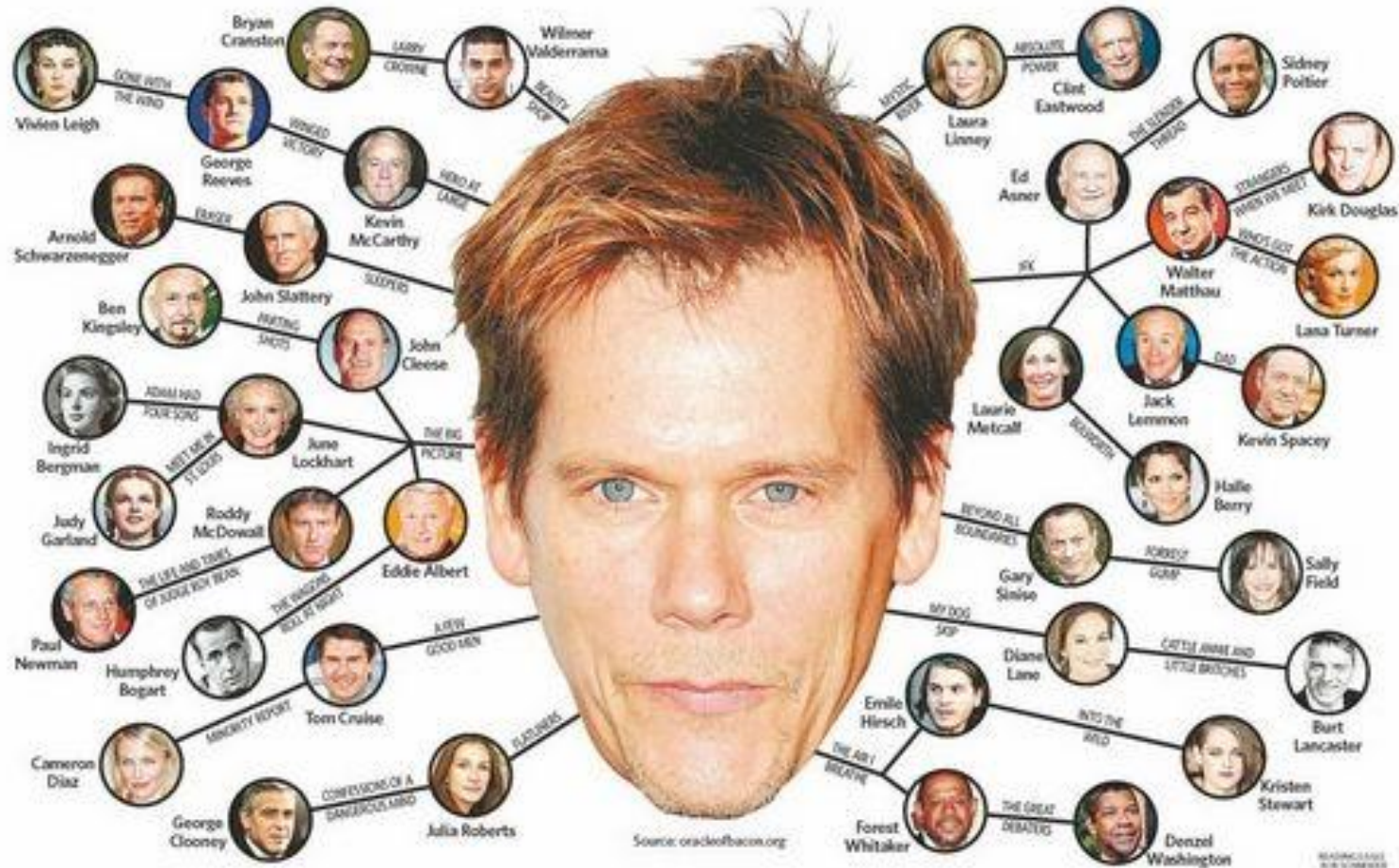
May 18, 2017

Network Analysis

- Looks at relationships between things, both mathematically and visually



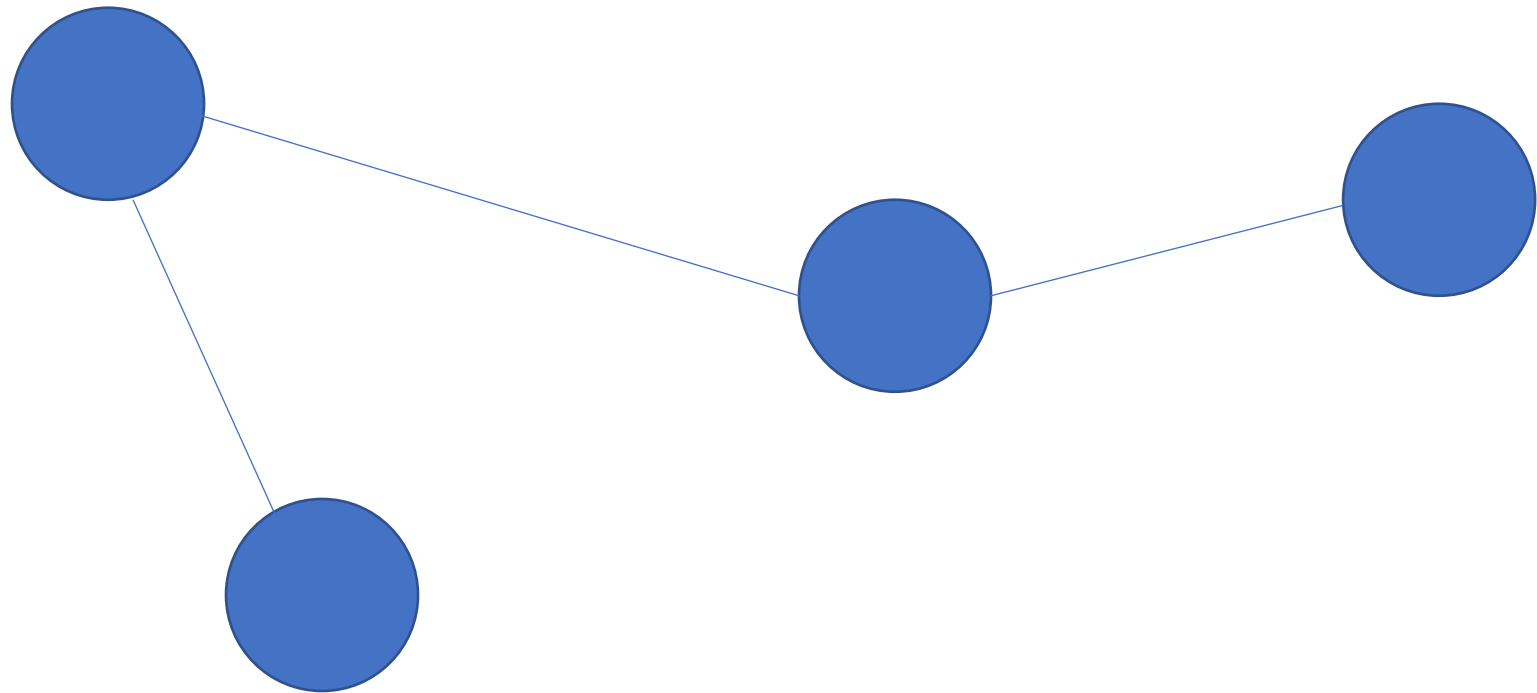
Six Degrees of Kevin Bacon



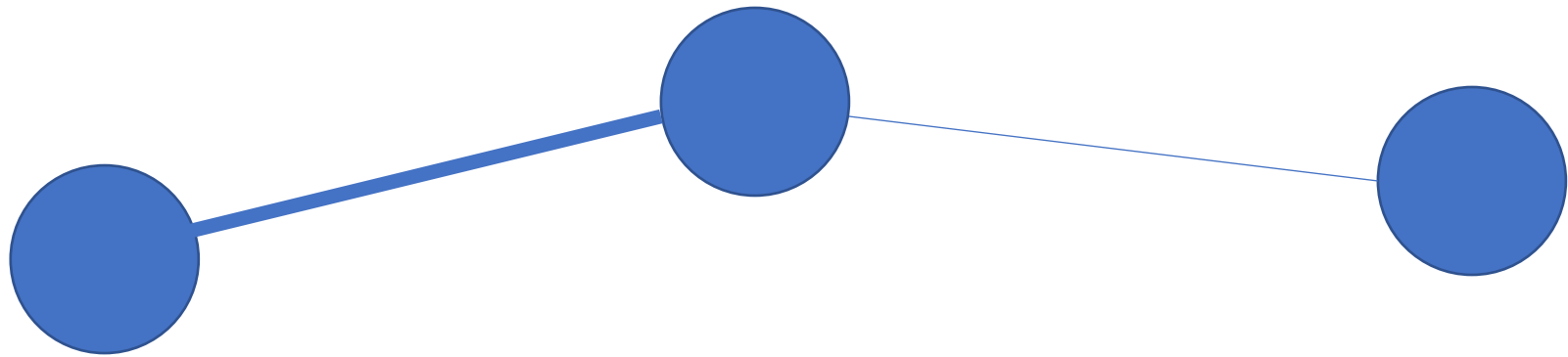
Some vocabulary....

- Nodes – people, things, ideas, places
- Edges – the connections between them
- Degree – how many other nodes connect to a given node?
- Centrality – what is the place of a given node in the larger network?

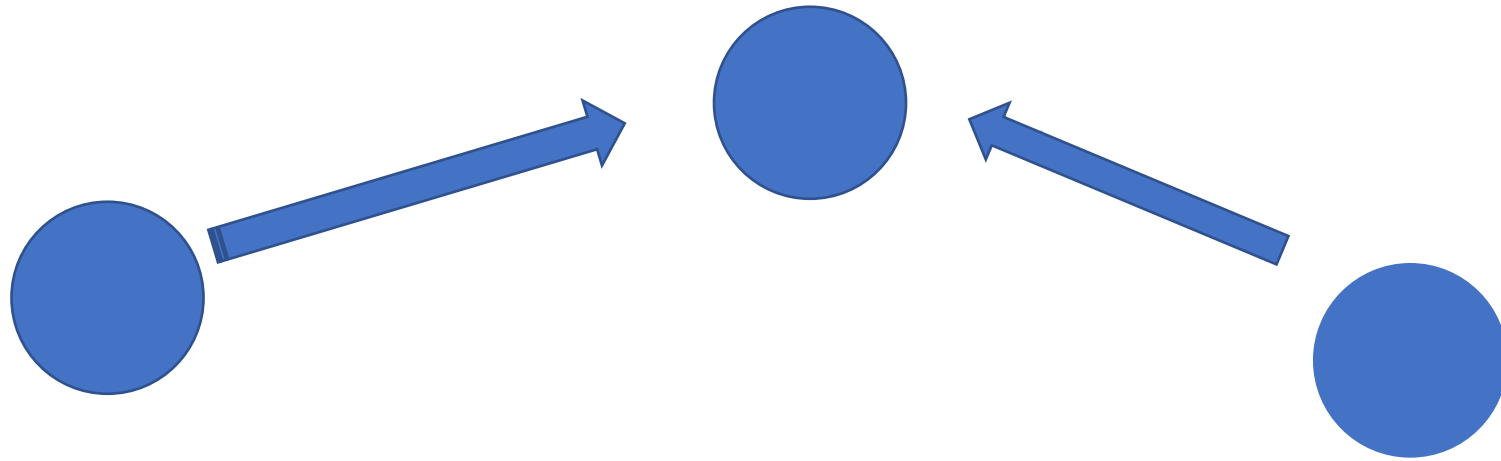
Nodes and Edges



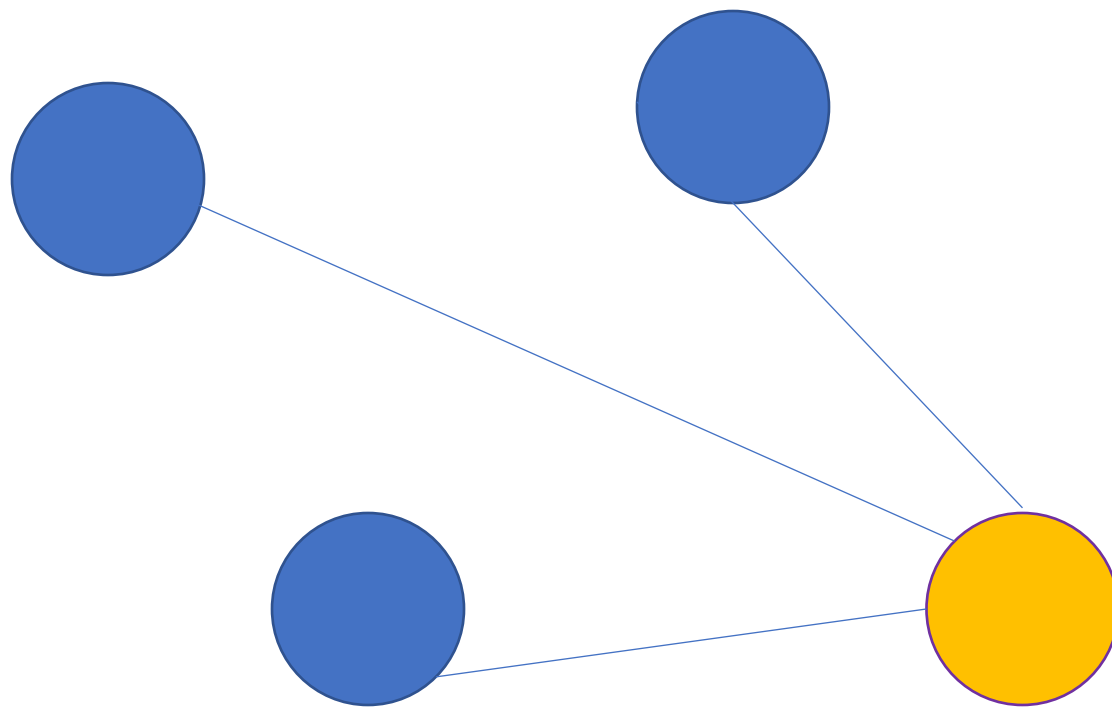
Edges can also have weight....



...or direction
///..



Single vs Multi Nodal



What can network analysis do?

- determine what the most connected nodes are (i.e. the most “important” in the network - if a given node were taken out, would the whole network collapse?)
- the degree of network centrality (is it a tightly focused network or a loose amalgamation?)

Ultimately, network analysis is all about relationships....

So what does it all mean?

- Relationships will be clearer
- See who the “gate keepers” are
- See how tightly connected a network is
- See which nodes share connections to other nodes (i.e. historical societies that share members or recipes that call for both egg whites and caraway)

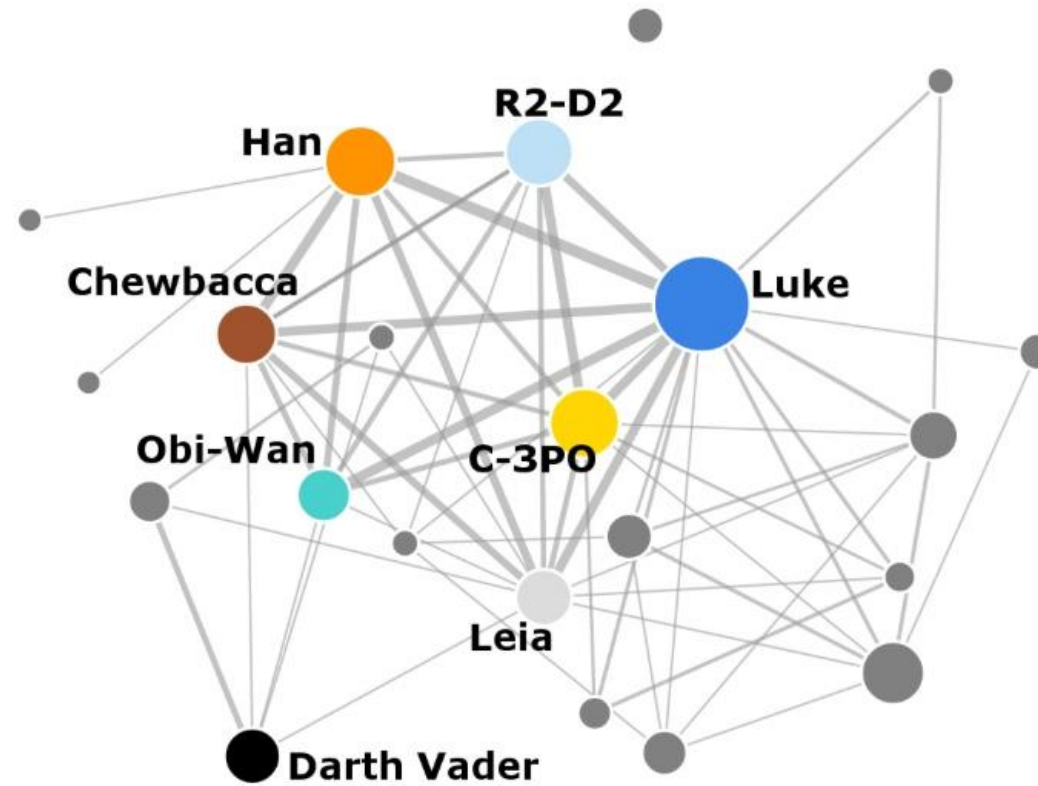
Network analysis takes all of that data and makes it easier to see, which leads to even more questions....

Some uses for network analysis

- ecological models
- modeling the spread of a disease
- determining key people or positions (i.e. research matchmaking)
- looking at relationships between ideas
- citation analysis

Example

Episode IV: A New Hope



- Evelina Gabasova (<http://evelinag.com/blog>)

Why use Gephi?

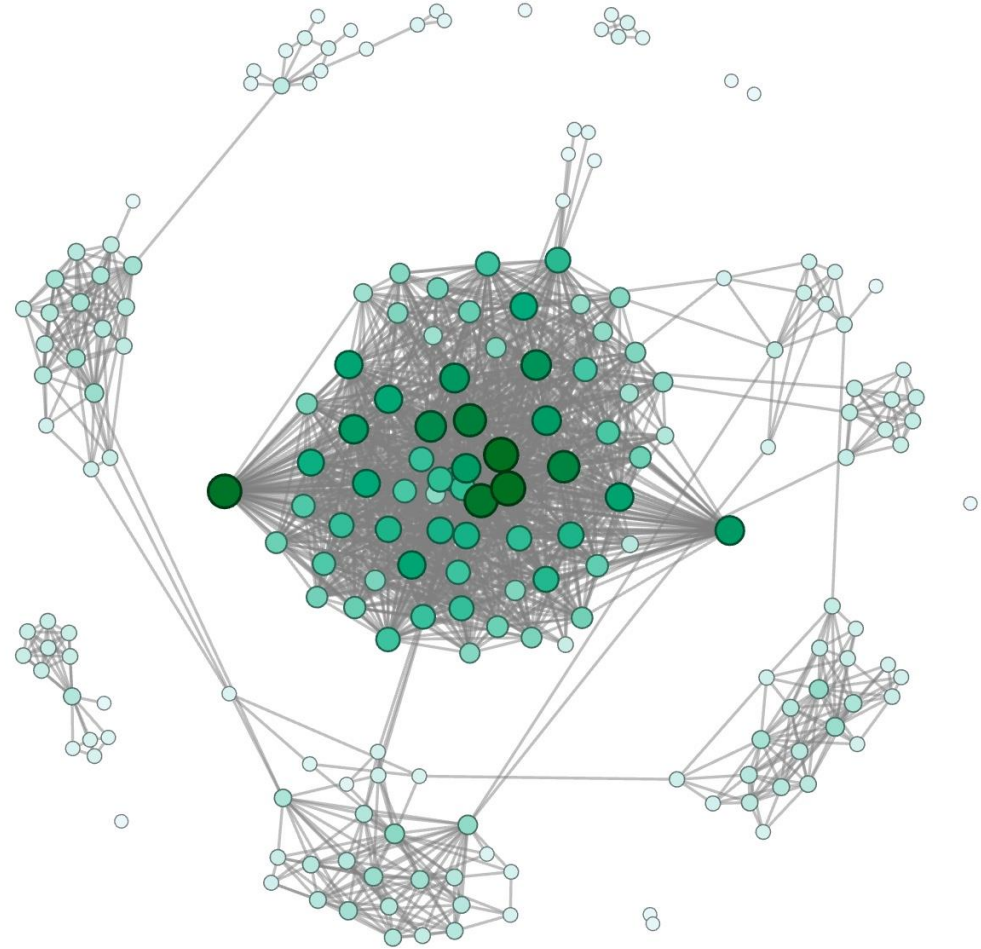
- Platform robust (PC, Mac, Linux)
- Open source
- Lets you do both the simple and complex
- Attractive visualizations
- No coding required

Limitations to Gephi

- Works best for single nodal/edge types
- Less flexible and nuanced than coding languages such as R
- Support is less robust (updates don't always happen frequently enough)

Example:

Facebook Network



Before you start...



What is your question?

Fussy Data (Data Prep)

Gephi wants

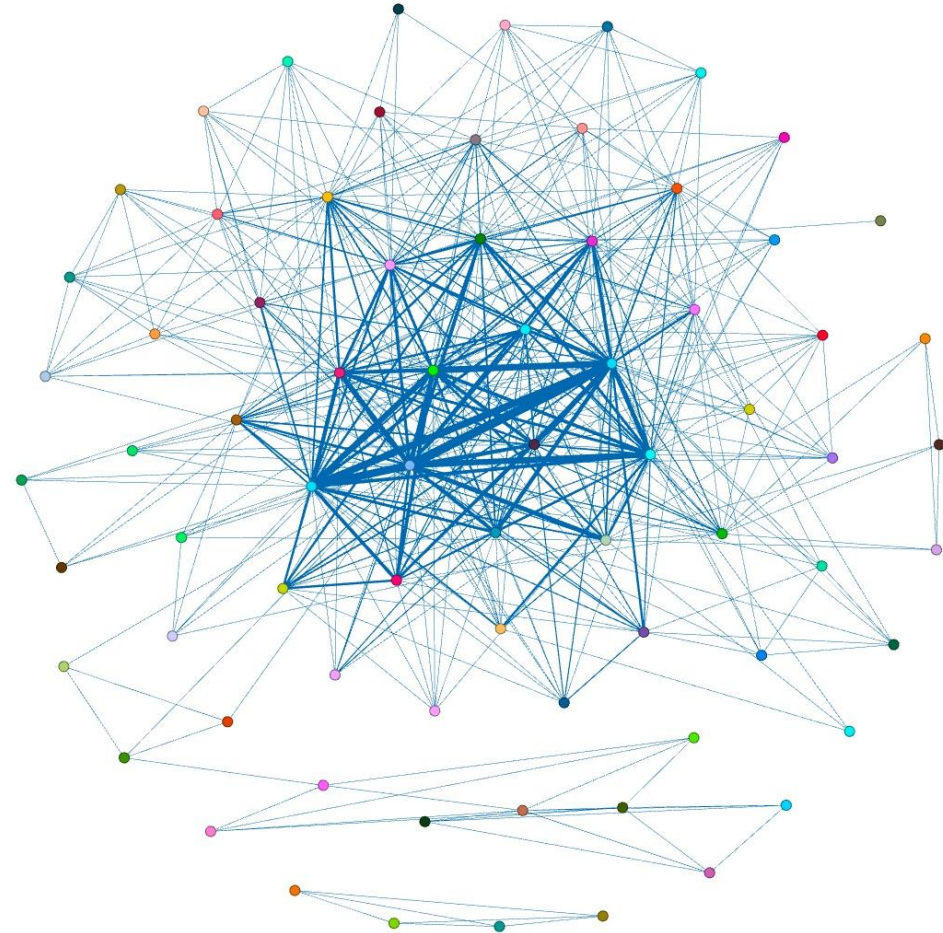
- A list of unique nodes with numerical identifiers (can be imported as a .csv file)
- A list of edges between those numerically identified nodes in the form of “source” and “target” (two columns), with optional “weight” and “direction” columns (also can be imported as a .csv file)
- Watch out – if you use Excel, Gephi may not recognize your file as .csv (in that case, just copy and paste your values into a new .csv file)

Recommended Tutorial:

- Data Preparation for Gephi: Step by Step (Song Chen)
<http://humnviz.blogs.bucknell.edu/files/2015/11/Data-Preparation-for-Gephi.pdf>

Example:

Fairfax Family
Cookbook
(George Mason
University Libraries)



Resources

- Data Preparation for Gephi: Step by Step (Song Chen)
<http://humnviz.blogs.bucknell.edu/files/2015/11/Data-Preparation-for-Gephi.pdf>
- The Complete n00b's Guide to Gephi (Brian Sarnacki)
<http://www.briansarnacki.com/gephi-tutorial/>
- Creating a Network Graph with Gephi (Miriam Posner)
<http://miriamposner.com/dh101f14/wp-content/uploads/2014/11/Creating-a-Network-Graph-with-Gephi.pdf>
- Demystifying Networks (Scott Weingart)
<http://www.scottbot.net/HIAL/index.html@p=6279.html>

Image Credits

- Social Network Analysis Example (Wikipedia)
https://en.wikipedia.org/wiki/Social_network_analysis
- Six Degrees of Kevin Bacon.
<http://www.readingeagle.com/life/article/six-degress-of-kevin-bacon-a-game-changer>
- Star Wars Social Network Graph
<http://evelinag.com/blog/2015/12-15-star-wars-social-network/#.WRd3TIKZM6U>
- Monty Python's Holy Grail Knights
<http://www.retroland.com/monty-python-and-the-holy-grail/>